Classification of Hyperspectral Image (HSI) using transfer learning and deep neural networks

Student Name: Smit Panchal

Student Id: 225817220

Student Name: Malay Patel

Student ID: 225816420

Abstract

Deep convolutional neural networks (CNNs) have demonstrated their exceptional performance in the classification of hyperspectral images (HSI). To identify and distinguish spatially connected material through thorough analysis, hyperspectral images provide a variety of spectral characteristics. A broad range of advanced classification approaches are feasible based on spectral data and spatial correlation. Modern, effective deep learning (DL) approaches have gained popularity due to advancements in computer technology, and their performance in a wide range of applications is encouraging. Deep learning techniques have been used successfully in different fields, particularly in data obtained by Earth Observer (EO) equipment for remote sensing. Given the wealth of information included in these sorts of photographs, hyperspectral imaging (HSI) is one of the primary subjects in remote sensing research which enables increased earth surface analysis and exploitation through the integration of rich spatial and spectral information. Given the enormous dimensions of the data and the limiting availability of training samples, HSI offers substantial challenges for supervised classification algorithms. Transfer learning architectures showing tremendous promise in the classification of HSI information have been designed recently to alleviate some restrictions. This paper describes an experiment of HSI classification using transfer learning. Experiment findings reveal that a restricted number of training sample models delivers improved accuracy.

1. Introduction

Different picture types, such as hyperspectral, multispectral, panchromatic, etc., are accessible for analysis in geographic information systems. This study classifies land use, mineral exploration, water pollution, and other phenomena using hyperspectral pictures. Hyperspectral Imaging is an analytical technique that focuses on spectroscopy. Various wavelengths and hundreds of images of the same space field are gathered. The human eye only has three colour receptors, red, green, and blue, the hyperspectral image checks the persistent light spectrum with a high wavelength resolution for each image pixel not just in visible, but also in near-infrared. The resulting information is referred to as a hyperspectral cube, in which the spatial scale and the third of its spectrum components are each represented

by two measurements. Multispectral imaging produces photos with up to 10 wider bands, whereas hyperspectral imaging creates images with hundreds of bands. One's ability to perceive the unseen increases with enhanced spectral information in HSI.

In recent years, supervised and unsupervised machine learning techniques have been used for the classification of HSI. Deep learning models have recently outperformed older machine learning methods in classification tasks. Convolution layers are used by the majority of classification deep neural networks to extract important characteristics from the data.

This paper use transfer learning methodology for classification. Transfer learning is a machine learning strategy that pretrains a model outside of the original domain to reduce the requirement for a large amount of labelled training data [1]. According to our method, the initial CNN model for training is a model that has been pretrained using the 1.2 million-image ImageNet dataset. It is important to note that ImageNet is the largest and most widely used dataset for picture classification and detection [2]. Pre-training is helpful for configuring our CNN model and avoiding problems caused by a lack of training data. For classification in transfer learning, we used pre-trained deep neural network models like VGG-16, ResNet-50, and also paper will use modified LinkNet architecture.

2. Related Work

Wang et al. [3] presented a salient band-selected methodology for hyperspectral image classification. In this methodology, convolution is performed to characterise bands into groups of bands, and from that, each group representative is selected. After that, it uses a manifold ranking strategy to select a salient band from selected representatives. The performance of this methodology depends on the quality of selected representatives and manifold. Some researchers use kernel-based methodology for the classification of HSI. In kernel-based classification methods, the HSI is translated to higher dimension space until it becomes linearly separable. After that, it is easy to classify images. Camps et al. [4] implemented a kernel-based classification methodology it has given significant results for classification. There are some difficulties in kernel methods such as sometimes it is hard to translate to a higher dimension and also kernel need to be chosen carefully because model performance depends on it.

A transfer learning strategy for HSI classification using a customized CNN model was introduced by Xin He et al. [5]. There are considerable differences between HSI and digital images. In order to fully utilize the potential of transfer learning, heterogeneous transfer learning based on CNN is introduced for hyperspectral image classification. In this, they have presented a technique that involves taking the weight from only the first seven layers of the VGG-16 Net and applying it to a specially created CNN model. The attention function is used in this study to reduce the detrimental effects of the difference between the digital image and HSI. Using 200

samples from the Indian Pines and Salinas Valley databases, they were able to obtain 87% and 93% testing accuracy in this case. The Markov random fields in combination with the deep convolution layer were used for classification of HSI [6]. This methodology has given 89% accuracy on the BigEarthNet dataset. H.Lee [7] build deeper and wider convolution neural networks compare to existing deep CNNs for HSI classification. The model is made of 9 layers and for the training of sparse samples, it used residual learning methodology[8]. It has given an accuracy of 93.61% on Indian pines and 95.07% on the Salinas dataset.

Different machine learning and deep learning models are developed and contrasted by M.E. Paoletti et al. [9] for the classification of hyperspectral images. Six distinct methodologies have been used, and they have also presented their results and a comparison of several models. This includes a thorough examination of the most recent state-of-the-art in deep learning for the classification of hyperspectral images, highlighting the benefits and drawbacks of the most popular classification models in the research. With the help of numerous frequently used hyperspectral image scenarios, they have provided a comparative study for each method outlined, providing a thorough overview of the investigated techniques.

Deep learning approaches have been popular since AlexNet's victory in the 2012 ImageNet classifications competition. AlexNet can find the best discriminating characteristics. Recently, a number of deep-learning methods for HSI classification have been developed. In an effort to classify HSI, Chen et al. [10] tried to combine a deep spatial domain technique based on auto-encoder (SAE) stacked with PCA-dependent dimensionality reductions and logistic regression classification.

3. Methodology

This portion will cover pre-processing of HSI datasets and details about different models that are used in this paper to perform HSI classification.

3.1 Pre-processing of dataset

The HSI are made of multiple bands but in this study, we are using deep learning models which take input image which has 3 channels, therefore, we are performing dimensionality reduction on the dataset. This paper uses principal component analysis (PCA) for dimensionality reduction. The PCA algorithm will reduce dimensionality by reducing spatial redundancy as well as it will maintain important spatial information [11].

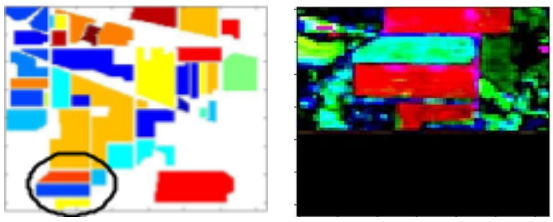

Figure 1: Ground truth image and pre-processed image

The figure 1 shows the similarity between the ground truth image and the pre-processed image. Here, the left side figure is one portion from the ground truth images from the dataset and the right side is the pre-processed image. The PCA is

performed on both Indian pines and salinas datasets.

3.2 Classifier deep learning models

This paper uses three different deep convolutional neural networks(CNNs) for the comparative study. For Indian pines dataset, models will take input of dimension 64*64*3 and in the case of salinas, it would be of 32*32*3 dimensions. For both databases, the output dimension will remain 16. The pre-trained model gives output features of different dimensions to adjust its paper using the attention mechanism [12].

3.2.1 Vgg-16

The full form of VGG is Visual Geometry Group. The VGG-16 is a 16 layers deep convolutional neural network [13]. The VGGNet was introduced by A. Zisserman and K. Simonyan from the University of Oxford. The model was submitted to the ILSVRC-2014 competition. It has given 92.7% accuracy on imageNet dataset.

Figure 3. explains the architecture of the VGG-16 model which consist of 13 convolution layer with a kernel size of 3 and 3 fully connected layers.

3.2.2 ResNet

Deep residual networks (ResNets), such as the popular ResNet-50 model, are another type of convolutional neural network architecture (CNN) that is 50 layers deep [14]. The model was submitted to ILSVRC-2015. Compared to VGGNets, ResNets are less complex since they have fewer filters. In this paper, we have used ResNet-50.

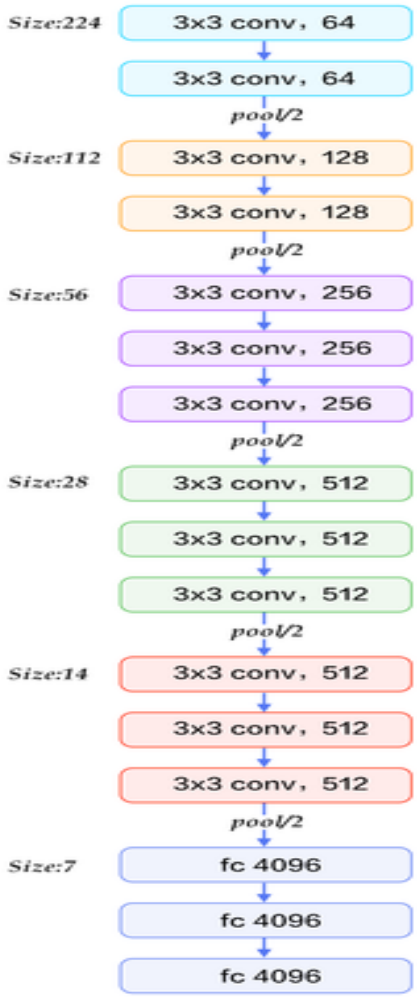

Figure 2. VGG-16 architecture

3.2.3 LinkNet

The LinkNet model focuses on real-time segmentation by utilizing the resources of an embedded system. The LinkNet uses 5.6 million parameters which is relatively very less compared to VGG16(138 million parameters) and AlexNet(61 million parameters) therefore its computational time is comparatively low. The architecture of LinkNet is made of a ladder of encoders and

decoders [15].

From the below figure, we can see that in LinkNet sampled feature maps of the encoders are summed up with unsampled feature maps of the decoders by this way spatial information lost during multiple upsampling operations of the encoder is recovered. The encoder part has four convolutional layers and the decoder part has a full-convolution layer between two convolutional layers.

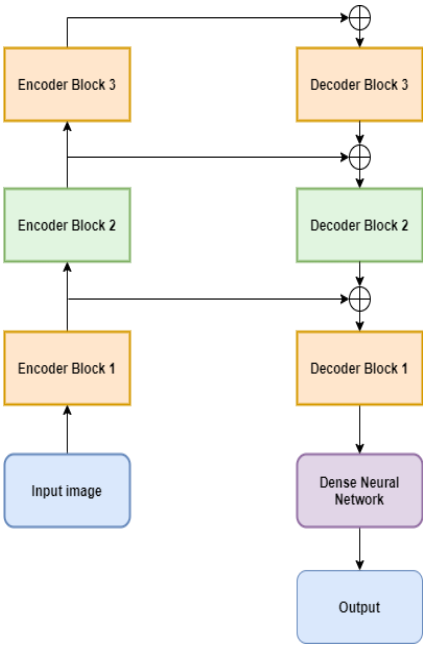

Figure 3: Linknet

In LinkNet batch normalization is used between each convolution layer and the ReLU layer is used for non-linearity. This paper uses modified LinkNet without pre-training to compare the result of it with pre-trained models.

4. Experiments and Results

This section will cover details about experiment constraints and results driven by different models.

4.1 Experiment

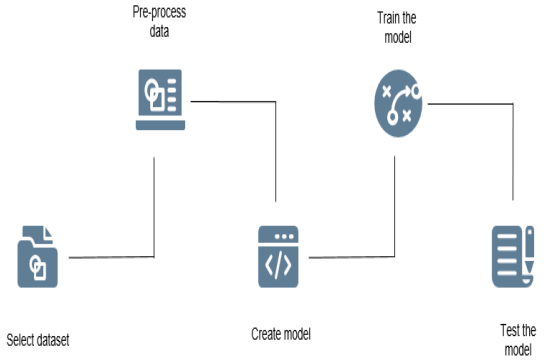

Figure 4 Implementation flow

Figure 4 describes the implementation flow in which first step we will select a dataset from Indian pines and salinas. The second step will perform PCA on the dataset and reduce the dimension of multiband spectral image to 3 channels [16]. After PCA, the standardization is performed using standardscaler. For analysis, we made function createImageCubes which will make small image cubes so that we can visualize how HSI looks after pre-processing. In the third step, the VGG-16, ResNet-50, and LinkNet models are created. In step 4, the training of the model is done using pre-processed data. Here, the Indian pines dataset is used with 3% training whereas Salinas dataset is used with 1% training. For both, the dataset 10% was used for validation and the remaining for testing.

4.2 Training Specification

In this paper for performance evolution precision, recall, and f1-score are used.

4.3 Results

In this paper for performance evolution precision, recall, and f1-score are used.

IndianPines									
	VGGNet			ResNet 50			LinkNet		
	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
Alfalfa	1.00	0.37	0.54	1.00	0.34	0.51	0.41	0.59	41
Corn-notill	0.95	0.86	0.91	0.92	0.94	0.93	0.94	0.95	0.95
Corn-mintill	0.89	0.94	0.92	0.90	0.94	0.92	0.92	0.92	0.92
Corn	0.95	0.95	0.95	0.93	1.00	0.96	0.98	0.87	0.92
Grass-pasture	0.99	0.78	0.87	0.99	0.74	0.85	0.98	0.8	0.88
Grass-trees	0.96	0.97	0.96	0.94	0.94	0.94	0.94	0.97	0.96
Grass-pasture-mowed	0.42	1.00	0.59	0.49	1.00	0.66	0.57	1	0.73
Hay-windrowed	0.99	1.00	1.00	1.00	1.00	1.00	1	1	1
Oats	0.50	0.18	0.26	1.00	0.29	0.45	1	0.35	0.52
Soybean-notill	0.84	0.96	0.89	0.93	0.96	0.94	0.89	0.91	0.9
Soybean-mintill	0.96	0.87	0.91	0.95	0.97	0.96	0.96	0.89	0.92
Soybean-clean	0.89	0.95	0.92	0.93	0.94	0.93	0.9	0.87	0.89
Wheat	0.93	0.91	0.92	0.88	0.91	0.90	0.93	0.98	0.96
Woods	1.00	0.97	0.98	0.99	0.97	0.98	0.97	0.97	0.97
Buildings-Grass-Trees-Drives	0.58	0.99	0.73	0.93	1.00	0.96	0.55	0.96	0.7
Stone-Steel-Towers	0.84	0.86	0.85	1.00	0.70	0.83	0.95	0.65	0.77
Accuracy	0.91			0.94			0.92		

Table1. Results of Indian pines dataset

Salinas									
	VGGNet			ResNet 50			LinkNet		
	Precision	Recall	F1 score	Precision	Recall	F1 score	Precision	Recall	F1 score
Brocoli_green_weeds_1	0.90	1.00	0.95	0.32	0.01	0.02	0.82	0.92	0.86
Brocoli_green_weeds_2	1.00	0.93	0.96	1.00	0.07	0.13	0.95	0.89	0.92
Fallow	0.85	0.97	0.90	0.20	0.39	0.26	1.00	0.99	0.99
Fallow_rough_plow	1.00	0.91	0.95	0.90	0.22	0.35	1.00	0.97	0.99
Fallow_smooth	0.96	0.98	0.97	0.69	0.16	0.26	0.99	1.00	0.99
Stubble	0.99	1.00	0.99	1.00	0.16	0.27	1.00	1.00	1.00
Celery	0.99	0.97	0.98	0.35	0.39	0.37	1.00	0.98	0.99
Grapes_untrained	0.97	0.98	0.98	0.31	0.11	0.16	0.94	0.94	0.94
Soil_vinyard_develop	0.98	0.99	0.98	0.30	0.05	0.09	0.99	1.00	1.00
Corn_senesced_green_weeds	0.99	0.97	0.98	0.00	0.00	0.00	0.97	1.00	0.99
Lettuce_romaine_4wk	1.00	0.95	0.97	0.99	0.07	0.13	0.99	0.93	0.96
Lettuce_romaine_5wk	0.93	0.93	0.93	0.10	0.19	0.13	0.97	0.99	0.98
Lettuce_romaine_6wk	1.00	0.87	0.93	0.03	1.00	0.05	1.00	0.95	0.97
Lettuce_romaine_7wk	0.99	0.95	0.97	0.01	0.00	0.00	0.99	0.98	0.98
Vinyard_untrained	0.97	0.97	0.97	0.11	0.00	0.00	0.90	0.90	0.90
Vinyard_vertical_trellis	1.00	1.00	1.00	0.00	0.00	0.00	0.98	0.99	0.98
Accuracy	0.97			0.12			0.96		

Table 2. Result of salinas dataset

5. Discussion on Results

Table 1 and Table 2 depict the obtained results. Looking at the Indian Pines dataset, it can be determined that the ResNet50 obtains the maximum accuracy with just 3% of labeled samples, achieving 94% accuracy. Then comes the VGG16 and LinkNet with 91% and 92% respectively. In this regard, it should be highlighted that the Indian Pines dataset is more complicated

than the Salinas sceneries as there are only a few samples available for some of the classes. Although these deep learning models are not specially built for HSI, they can produce fascinating results.

In the case of Salinas dataset, we were able to obtain the best result with VGG16 model and were able to achieve 97% accuracy, however, the LinkNet showed similar results and achieved 96% accuracy. While the ResNet50 model which showed better results for the Indian Pines dataset was not able to produce good results using this dataset.

The VGG-16 gives the best result for both datasets. In the case of ResNet-50 model shows overfitting for the salinas dataset because it is a very deep neural network. The modified LinkNet gives good results for both datasets considering it is not pre-trained on ImageNet dataset.

6. Conclusion

Deep CNN neural networks have shown good results in the field of HSI classification compare to previous machine learning techniques such as Support Vector Machine and Naive Bayes. The PCA captures important features of the HSI which makes it easier for models to train on the dataset.The pre-trained models are able to capture the details from both datasets and this hybrid methodology to combine deep CNN with transfer learning works well for HSI classification.The addition of layers to pre-trained models converts pre-trained model output features to desired dataset output features. In this paper dataset size for different classes is less so that for some classes the f1-score is comparatively low. To mitigate it we have to use a dataset that has more data samples so that model can equally train on all classes. In future work, performance for the HSI dataset can be increased using more advanced pre-trained models such as Inception-V3 and PNASNet-5.

References

[1] S. J. Pan and Q. Yang, “A survey on transfer learning,” IEEE Trans.Knowl. Data Eng., vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[2] J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248-255.

[3] Q. Wang, J. Lin, and Y. Yuan, “Salient band selection for hyperspectral image classification via manifold ranking,” IEEE transactions on neural networks and learning systems, vol. 27, no. 6, pp. 1279–1289, 2016.

[4] G. Camps-Valls and L. Bruzzone, “Kernel-based methods for hyperspectral image classification,” IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 6, pp. 1351–1362, 2005.

[5] X. He, Y. Chen and P. Ghamisi, "Heterogeneous Transfer Learning for Hyperspectral Image Classification Based on Convolutional Neural Network," in IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 5, pp. 3246-3263, May 2020, DOI: 10.1109/TGRS.2019.2951445.

[6] X. Cao, F. Zhou, L. Xu, D. Meng, Z.

Xu, and J. Paisley, “Hyperspectral image segmentation with Markov random fields and a convolutional neural network,” arXiv preprint arXiv:1705.00727, 2017.

[7] H. Lee and H. Kwon, "Going Deeper With Contextual CNN for Hyperspectral Image Classification," in *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4843-4855, Oct. 2017, DOI: 10.1109/TIP.2017.2725580.

[8] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[9] M.E. Paoletti, J.M. Haut, J. Plaza, A. Plaza, Deep learning classifiers for hyperspectral imaging: A review, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 158,2019, Pages 279-317, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2019.09.006.

[10] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, “Deep learning-based classification of hyperspectral data,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 7, no. 6, pp. 2094–2107, Jun. 2014.

[11] Huang, Shiqi, et al. "Multi-scale guided feature extraction and classification algorithm for hyperspectral images." Scientific Reports 11.1 (2021): 1-13.

[12] K. Xu et al., “Show, attend and tell: Neural image caption generation with visual attention,” in Proc. ICML, 2015, pp. 2048–2057.

[13] H. Wang, "Garbage Recognition and Classification System Based on Convolutional Neural Network VGG16," 2020 3rd International Conference on Advanced Electronic Materials, Computers, and Software Engineering (AEMCSE), 2020, pp. 252-255, DOI: 10.1109/AEMCSE50948.2020.00061.

[14] Mukti, Ishrat Zahan, and Dipayan Biswas. "Transfer learning based plant diseases detection using ResNet50." 2019 4th International conference on electrical information and communication technology (EICT). IEEE, 2019.

[15] Guo, Xiaofei, et al. "A LinkNet Deep Learning Model for Daytime Sea Fog Detection." Remote Sensing 13.24 (2021): 5163.

[16] Patel, U., Patel, S., Kathiria, P. (2022). Hyperspectral Image Classification Using Transfer Learning. In: Sharma, H., Shrivastava, V., Kumari Bharti, K., Wang, L. (eds) Communication and Intelligent Systems. Lecture Notes in Networks and Systems, vol 461. Springer, Singapore. https://doi.org/10.1007/978-981-19-2130-8_43